**ECE 509: Convex Optimization** **Rutgers: Spring 2024**

# Homework #7

April 11, 2024

*Name: Ravi Raghavan* *Extension: No*

---

**Problem 1.** *Voronoi description of halfspace.* Let $a$ and $b$ be distinct points in $\mathbb{R}^n$. Show that the set of all points that are closer (in Euclidean norm) to $a$ than $b$, *i.e.*, $\{x : ||x - a||_2 \le ||x - b||_2\}$, is a halfspace. Describe it explicitly as an inequality of the form $c^T x \le d$. Draw a picture.

**Solution.** We can see that the following two sets are equivalent:

$S_1 = S_2$ where $S_1 = \{x : ||x - a||_2 \le ||x - b||_2\}$ and $S_2 = \{x : ||x - a||_2^2 \le ||x - b||_2^2\}$

Let's work with $S_2$ since it will be a lot easier

$S_2 = \{x : ||x - a||_2^2 \le ||x - b||_2^2\}$
$S_2 = \{x : ||x||_2^2 - 2 < x, a > + ||a||_2^2 \le ||x||_2^2 - 2 < x, b > + ||b||_2^2\}$
$S_2 = \{x : -2 < x, a > + ||a||_2^2 \le -2 < x, b > + ||b||_2^2\}$
$S_2 = \{x : 2 < x, b - a > \ \le ||b||_2^2 - ||a||_2^2\}$
$S_2 = \{x : 2(b - a)^T x \ \ \le ||b||_2^2 - ||a||_2^2\}$
$S_2 = \{x : (b - a)^T x \ \ \le 0.5(||b||_2^2 - ||a||_2^2)\}$

If we set $c = b - a$ and $d = 0.5(||b||_2^2 - ||a||_2^2)$, we can express $S_2$ as follows:
$S_2 = \{x : c^T x \ \ \le d\}$
This is a closed half-space

---

**Problem 2.** Which of the following sets $S$ are polyhedra? If possible, express $S$ in the form $S = \{x | Ax \preceq b, Fx = g\}$

(b) Yes $S$ is a polyhedra.

Let $M_1 = [a_1, a_2, ..., a_n] \in \mathbb{R}^{1xn}$ and let $M_2 = [a_1^2, a_2^2, ..., a_n^2] \in \mathbb{R}^{1xn}$

Let $F$ be the vertical concatentation of $1^T$, $M_1$, and $M_2$. Let $g$ be $[1, b_1, b_2]^T$

Let $A = -I$

We can express $S$, via compact notation, as $S = \{x | Ax \preceq 0, Fx = g\}$

Note: Just to confirm, $\preceq$ when applied to 2 vectors means component-wise $\le$

(c) $S$ is NOT a Polyhedra. Let's analyze why this is the case.

First we must look at the statement $x^T y \le 1$ for all $||y||_2 = 1$. Our goal is to take all these inequalities and express them in the form $Ax \preceq b$, where $A$ is a matrix, $\preceq$ when applied to 2 vectors means component-wise $\le$, and $b$ is a vector. So, given a $y$ where $||y||_2 = 1$, we can put this $y$ in a row of $A$. $Ax \preceq 1$, where each row of $A$ is a feasible value of $y$, would capture the statement $x^T y \le 1$ for all $||y||_2 = 1$.

However, the number of vectors $y$ where $||y||_2 = 1$ is infinite. Hence, our matrix $A$ would have to have infinite rows which is clearly not possible.

Another way of looking at it is like this:
$S$ is the intersection of an infinite number of halfspaces. As per definition, polyhedra are the intersection of a FINITE number of halfspaces.

We have proved why $S$ is NOT a polyhedra.

**Problem 3.** *Hyperbolic sets.* Show that the *hyperbolic set* is $\{x \in \mathbb{R}^2_+ : x_1 x_2 \geq 1\}$ is convex. As a generalization, show that $\{x \in \mathbb{R}^2_+ : \prod_{i=1}^n x_i \geq 1\}$ is convex. *Hint.* If $a, b \geq 0$ and $0 \leq \theta \leq 1$, then $a^\theta b^{1-\theta} \leq \theta a + (1-\theta) b$

**Solution.** Let's start by proving that the *hyperbolic set* is $S = \{x \in \mathbb{R}^2_+ : x_1 x_2 \geq 1\}$ is convex.
Let's have two vectors $j$ and $k$ that are in $S$. Let the elements of $j$ be $j_1, j_2$. Let the elements of $k$ be $k_1, k_2$

We want to prove that $\theta j + (1-\theta)k \in S$.

Since we know that $j \in S$ and $k \in S$, we can state the following:

- $j_1 \geq 0, j_2 \geq 0$
- $j_1 j_2 \geq 1$
- $k_1 \geq 0, k_2 \geq 0$
- $k_1 k_2 \geq 1$

For $0 \leq \theta \leq 1$, since $j_1 \geq 0, j_2 \geq 0$ and $k_1 \geq 0, k_2 \geq 0$, we can see that $0 \leq j_1^\theta k_1^{1-\theta} \leq \theta j_1 + (1-\theta)k_1$ and $0 \leq j_2^\theta k_2^{1-\theta} \leq \theta j_2 + (1-\theta)k_2$.
$j_1^\theta k_1^{1-\theta} j_2^\theta k_2^{1-\theta} = (j_1 j_2)^\theta (k_1 k_2)^{1-\theta}$

Since $j_1 j_2 \geq 1$ and $k_1 k_2 \geq 1$, we can say that:

$j_1^\theta k_1^{1-\theta} j_2^\theta k_2^{1-\theta} = (j_1 j_2)^\theta (k_1 k_2)^{1-\theta} \geq 1$

Since $j_1^\theta k_1^{1-\theta} \leq \theta j_1 + (1-\theta)k_1$ and $j_2^\theta k_2^{1-\theta} \leq \theta j_2 + (1-\theta)k_2$

$1 \leq j_1^\theta k_1^{1-\theta} j_2^\theta k_2^{1-\theta} \leq (\theta j_1 + (1-\theta)k_1)(\theta j_2 + (1-\theta)k_2)$

Furthermore, since $j_1 \geq 0, j_2 \geq 0$ and $k_1 \geq 0, k_2 \geq 0$, we know that $\theta j_1 + (1-\theta)k_1 \geq 0$ and $\theta j_2 + (1-\theta)k_2 \geq 0$

We have finished proving that $\theta j + (1-\theta)k \in S$.

Let $S$ be $\{x \in \mathbb{R}^2_+ : \prod_{i=1}^n x_i \geq 1\}$. Let's have two vectors $j$ and $k$ that are in $S$. Let the elements of $j$ be $j_1, j_2, ...., j_n$. Let the elements of $k$ be $k_1, k_2, ..., k_n$

We want to prove that $\theta j + (1-\theta)k \in S$.

Since we know that $j \in S$ and $k \in S$, we can state the following:

- $j_1 \geq 0, j_2 \geq 0, ...., j_n \geq 0$
- $j_1 j_2 j_3 ... j_n \geq 1$
- $k_1 \geq 0, k_2 \geq 0, ...., k_n \geq 0$
- $k_1 k_2 k_3 ... k_n \geq 1$

For $i \in [1, n]$, $j_i, k_i \geq 0$ and $0 \leq \theta \leq 1$, we can see that $0 \leq j_i^\theta k_i^{1-\theta} \leq \theta j_i + (1-\theta)k_i$.

We can also see that:

$\prod_{i=1}^{n} j_i^{\theta} k_i^{1-\theta} = (\prod_{i=1}^{n} j_i)^{\theta} (\prod_{i=1}^{n} k_i)^{1-\theta}$

Since $j_1 j_2 j_3 ... j_n \geq 1$ and $k_1 k_2 k_3 ... k_n \geq 1$, we can say that:

$\prod_{i=1}^{n} j_i^{\theta} k_i^{1-\theta} = (\prod_{i=1}^{n} j_i)^{\theta} (\prod_{i=1}^{n} k_i)^{1-\theta} \geq 1$

Since $j_i^{\theta} k_i^{1-\theta} \leq \theta j_i + (1-\theta) k_i$,

$1 \leq \prod_{i=1}^{n} j_i^{\theta} k_i^{1-\theta} \leq \prod_{i=1}^{n} \theta j_i + (1-\theta) k_i$

Furthermore, since $j_1 \geq 0, j_2 \geq 0, ...., j_n \geq 0$ and $k_1 \geq 0, k_2 \geq 0, ...., k_n \geq 0$, we know that $\theta j_i + (1-\theta) k_i \geq 0$ as well.

We have shown that $\theta j_i + (1-\theta) k_i \in S$ and that $S$ is a convex set!

---

**Problem 4.** Problem 2.16:
Show that if $S_1$ and $S_2$ are convex sets in $\mathbb{R}^{m+n}$, then so is their partial sum
$S = \{(x, y_1 + y_2) : x \in \mathbb{R}^m, y_1, y_2 \in \mathbb{R}^n, (x, y_1) \in S_1, (x, y_2) \in S_2\}$

**Solution.** Let's say that we have two points in $S$, namely $(x_1, y_{11} + y_{12})$ and $(x_2, y_{21} + y_{22})$. To prove that $S$ is convex, we need to show that $\theta(x_1, y_{11} + y_{12}) + (1-\theta)(x_2, y_{21} + y_{22})$ is in $S$.

Based on the definition of $S$, we can see the following:

- $(x_1, y_{11}) \in S_1$
- $(x_1, y_{12}) \in S_2$
- $(x_2, y_{21}) \in S_1$
- $(x_2, y_{22}) \in S_2$

Since $S_1$ and $S_2$ are convex, based on the definition of convex sets, we can see that:

- $\theta(x_1, y_{11}) + (1-\theta)(x_2, y_{21}) \in S_1$
  $(\theta x_1 + (1-\theta)x_2, \theta y_{11} + (1-\theta) y_{21}) \in S_1$
- $\theta(x_1, y_{12}) + (1-\theta)(x_2, y_{22}) \in S_2$
  $(\theta x_1 + (1-\theta)x_2, \theta y_{12} + (1-\theta) y_{22}) \in S_1$

By definition of Set $S$, we can see that:

$(\theta x_1 + (1-\theta)x_2, \theta y_{11} + (1-\theta) y_{21} + \theta y_{12} + (1-\theta) y_{22}) \in S$

$(\theta x_1 + (1-\theta)x_2, \theta(y_{11} + y_{12}) + (1-\theta)(y_{21} + y_{22})) \in S$

$\theta(x_1, y_{11} + y_{12}) + (1-\theta)(x_2, y_{21} + y_{22})$ is in $S$.

---

**Problem 5** (Problem 2.19(a)). *Linear-fractional functions and convex sets.* Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be the linear-fractional function

$$f(x) = (Ax + b)/(c^T x + d), \;\; dom f = \{x | c^T x + d > 0\} \tag{1}$$

In this problem, we study the inverse image of a convex set $C$ under $f$, *i.e.,*

$$f^{-1}(C) = \{x \in dom f : f(x) \in C\} \tag{2}$$

For each of the following sets $C \subseteq \mathbb{R}^n$, give a simple description of $f^{-1}(C)$

**Solution.** Let's look at the halfspace $C = \{y : g^T y \leq h\}$(with $g \neq 0$).

$f^{-1}(C) = \{x : g^T((Ax + b)/(c^T x + d)) \leq h, c^T x + d > 0\}$

Since $c^T x + d > 0$
$f^{-1}(C) = \{x : g^T(Ax + b) \leq h(c^T x + d), c^T x + d > 0\}$
$f^{-1}(C) = \{x : g^T Ax + g^T b \leq hc^T x + hd, c^T x + d > 0\}$
$f^{-1}(C) = \{x : (g^T A - hc^T)x \leq hd - g^T b, c^T x + d > 0\}$
$f^{-1}(C) = \{x : (A^T g - ch^T)^T x \leq hd - g^T b, c^T x + d > 0\}$

Let's call a new vector $p^T = (A^T g - ch^T)^T$ and $q = hd - g^T b$
$f^{-1}(C) = \{x : p^T x \leq q, c^T x + d > 0\}$

We can see that $f^{-1}(C)$ is just the intersection of a halfspace and the domain of f!

---

**Problem 6** (Problem 3.17). Suppose $p < 1, p \neq 0$. Show that the function

$$f(x) = (\sum_{i=1}^{n} x_i^p)^{\frac{1}{p}} \tag{3}$$

with $dom f = \mathbb{R}^n_{++}$ is concave. This includes as special cases $f(x) = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$ and the *harmonic mean* $f(x) = (\sum_{i=1}^n \frac{1}{x_i})^{-1}$. *Hint.* Adapt the proofs for the log-sum-exp function and the geometric mean in 3.1.5

**Solution.** Gradient: $\frac{\partial f}{\partial x_k} = (\sum_{i=1}^n x_i^p)^{\frac{1}{p}-1} x_k^{p-1}$

Jacobian:
$\frac{\partial^2 f}{\partial^2 x_k} = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2}[x_k^{2p-2} - x_k^{p-2} \sum_{i=1}^n x_i^p]$
$\frac{\partial^2 f}{\partial x_k \partial x_l} = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2} x_l^{p-1} x_k^{p-1}$

It is clear that we can express $\nabla^2 f(x)$ as follows:
$\nabla^2 f(x) = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2}[qq^T - (\sum_{i=1}^n x_i^p) diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})]$ where $q_i = x_i^{p-1}$

For concavity, we need $v^T \nabla^2 f(x) v \leq 0$ to hold true for all $x \in \mathbb{R}^n_{++}$

$v^T \nabla^2 f(x) v = (1-p)(\sum_{i=1}^n x_i^p)^{-\frac{1}{p}}[v^T qq^T v (\sum_{i=1}^n x_i^p)^{\frac{2}{p}-2} - (\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1} v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v]$

$v^T \nabla^2 f(x) v = (1-p)(\sum_{i=1}^n x_i^p)^{-\frac{1}{p}}[(\sum_{i=1}^n x_i^{p-1} v_i)^2((\sum_{i=1}^n x_i^p)^{\frac{1}{p}-1})^2 - (\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1} v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v]$

$v^T \nabla^2 f(x) v = (1-p)(\sum_{i=1}^n x_i^p)^{-\frac{1}{p}}[(\sum_{i=1}^n ((\sum_{i=1}^n x_i^p)^{\frac{1}{p}-1}) x_i^{p-1} v_i)^2 - \sum_{i=1}^n x_i^{p-2} v_i^2 ((\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1})]$

The Cauchy Schwartz Inequality tell us that $(a^Ta)(b^Tb) \geq (a^Tb)^2$
If we set $a_i = x_i^{\frac{p}{2}}((\sum_{i=1}^n x_i^p)^{\frac{-1}{2}})$ and we set $b_i = x_i^{\frac{p-2}{2}} v_i((\sum_{i=1}^n x_i^p)^{\frac{1}{p}-\frac{1}{2}})$, substituting this into the Cauchy Schwartz inequality allows us to see that

$$(\sum_{i=1}^n x_i^p)((\sum_{i=1}^n x_i^p)^{-1})(\sum_{i=1}^n x_i^{p-2}v_i)((\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1}) \geq (\sum_{i=1}^n x_i^{p-1}v_i)^2((\sum_{i=1}^n x_i^p)^{\frac{2}{p}-2})$$

Simplication of LHS[Left Hand Side]:
$(\sum_{i=1}^n x_i^p)((\sum_{i=1}^n x_i^p)^{-1})(\sum_{i=1}^n x_i^{p-2}v_i)((\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1}) = (\sum_{i=1}^n x_i^{p-2}v_i)((\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1})$

$= (\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1} v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v$
Simplication of RHS[Right Hand Side]:
$v^T qq^T v(\sum_{i=1}^n x_i^p)^{\frac{2}{p}-2}$

Since LHS $\geq$ RHS, we can state that:

$$(\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1} v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v \geq v^T qq^T v(\sum_{i=1}^n x_i^p)^{\frac{2}{p}-2}$$
$$v^T qq^T v(\sum_{i=1}^n x_i^p)^{\frac{2}{p}-2} - (\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1} v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v \leq 0$$

Since $p < 1$ and $x_i \in \mathbb{R}_{++}$, $(1-p)(\sum_{i=1}^n x_i^p)^{\frac{-1}{p}}$ is a positive constant.

$$v^T \nabla^2 f(x) v = (1-p)(\sum_{i=1}^n x_i^p)^{-\frac{1}{p}}[v^T qq^T v(\sum_{i=1}^n x_i^p)^{\frac{2}{p}-2} - (\sum_{i=1}^n x_i^p)^{\frac{2}{p}-1} v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v] \leq 0$$

Hence, since $v^T \nabla^2 f(x) v \leq 0$, we have proven concavity!

---

**Problem 7.** Adapt the proof of concavity of the log-determinant function in §3.1.5 to show the following

(a) $f(X) = tr(X^{-1})$ is convex on $dom f = \mathbb{S}_{++}^n$

**Solution.** To verify convexity, we can consider an arbitrary line given by $X = Z + tV$ where $Z \in \mathbb{S}_{++}^n$ and $V \in \mathbb{S}^n$

$g(t) = f(Z + tV) = tr((Z + tV)^{-1})$
$= tr((Z^{\frac{1}{2}}(I + tZ^{\frac{-1}{2}} V Z^{\frac{-1}{2}})Z^{\frac{1}{2}})^{-1})$
$= tr((Z^{\frac{-1}{2}}(I + tZ^{\frac{-1}{2}} V Z^{\frac{-1}{2}})^{-1} Z^{\frac{-1}{2}}))$

It is a well known fact that $tr(ABC) = tr(BCA) = tr(CAB)$

Hence, we can continue our simplification
$= tr((Z^{\frac{-1}{2}} Z^{\frac{-1}{2}}(I + tZ^{\frac{-1}{2}} V Z^{\frac{-1}{2}})^{-1}))$

$= tr((Z^{-1}(I + tZ^{\frac{-1}{2}} V Z^{\frac{-1}{2}})^{-1}))$

Let us represent $Z^{\frac{-1}{2}} V Z^{\frac{-1}{2}}$ as an Eigenvalue Decomposition of $QDQ^T$

$= tr((Z^{-1}(I + tQDQ^T)^{-1}))$
$= tr((Z^{-1}(QIQ^T + tQDQ^T)^{-1}))$
$= tr((Z^{-1}(Q(I + tD)Q^T)^{-1}))$

$= tr((Z^{-1}Q(I+tD)^{-1}Q^T))$

Again, we can use the cyclic property of the trace:
$= tr((Q^T Z^{-1}Q(I+tD)^{-1}))$

$= \sum_{i=1}^{n}(Q^T Z^{-1}Q)_{ii}(1+t\lambda_i)^{-1}$

Note: $\lambda_i$ are eigenvalues of $Z^{\frac{-1}{2}}VZ^{\frac{-1}{2}}$

Since $Z$ is a positive definite, symmetric matrix, we know that $Z^{-1}$ is a positive definite, symmetric matrix as well. Hence, $(Q^T Z^{-1}Q)_{ii}$ is always positive.

Let's now look at $(1+t\lambda_i)^{-1}$. Let's take the second derivative with respect to $t$. This second derivative is equal to $\frac{2\lambda_i^2}{(1+t\lambda_i)^3}$. The numerator is clearly positive.

We know that $Z + tV \in \mathbb{S}_{++}^n$
This means that $Z^{\frac{1}{2}}(I + tZ^{\frac{-1}{2}}VZ^{\frac{-1}{2}})Z^{\frac{1}{2}} \in \mathbb{S}_{++}^n$
This means that $(I + tZ^{\frac{-1}{2}}VZ^{\frac{-1}{2}}) \in \mathbb{S}_{++}^n$ must be True

This means that $1 + t\lambda_i > 0$ must be true for all $i$!. Hence, the denominator of $\frac{2\lambda_i^2}{(1+t\lambda_i)^3}$ is positive as well.

We have shown that the second derivative of $(1+t\lambda_i)^{-1}$ is positive over the values of $t$ that are in the domain of $g$ which means that $(1+t\lambda_i)^{-1}$ is a convex function over the domain of $g$.

$\sum_{i=1}^{n}(Q^T Z^{-1}Q)_{ii}(1+t\lambda_i)^{-1}$ is a non-negative weighted sum of convex functions which is convex!

We have completed the proof!

---

**Problem 8.** *Nonnegative weighted sums and integrals*

(a) Show that $f(x) = \sum_{i=1}^{r} \alpha_i x_{[i]}$ is a convex function of $x$, where $\alpha_1 \geq \alpha_2 \geq \dots \alpha_r \geq 0$, and $x_{[i]}$ denotes the $ith$ largest component of $x$. (You can use the fact that $f(x) = \sum_{i=1}^{r} x_{[i]}$ is convex on $\mathbb{R}^n$)

**Solution.** $f(x) = \alpha_r(\sum_{i=1}^{r} x_{[i]}) + (\alpha_{r-1} - \alpha_r)(\sum_{i=1}^{r-1} x_{[i]}) + (\alpha_{r-2} - \alpha_{r-1})(\sum_{i=1}^{r-2} x_{[i]}) + \cdots + (\alpha_1 - \alpha_2)(x_{[1]})$
Since $a_i > a_{i+1}$, $a_i - a_{i+1} > 0$.

We know that a non-negative combination of convex functions is convex. Hence, we have completed the proof

---

(b) Let $T(x, w)$ denote the trigonometric polynomial

$$T(x, w) = x_1 + x_2 \cos w + x_3 \cos 2w + \cdots + x_n \cos{(n-1)w} \tag{4}$$

Show that the function

$$f(x) = -\int_0^{2\pi} \log T(x, w)\, dw \tag{5}$$

is convex on $\{x \in \mathbb{R}^n : T(x, w) > 0, 0 \leq w \leq 2\pi\}$

**Solution.** Let $g(x,w) = -\log T(x,w) = -\log(x_1 + x_2 \cos w + x_3 \cos 2w + \cdots + x_n \cos{(n-1)w})$

Let's prove that $g(x,w)$ is convex in $x$ when we fix $w$

$h(t) = g(z+tv,w) = -\log T(z+tv,w) = -\log(z_1 + tv_1 + (\cos w)(z_2+tv_2) + (\cos 2w)(z_3+tv_3) + \cdots + (\cos{(n-1)w})(z_n+tv_n))$

$h'(t) = -\frac{v_1 + v_2\cos w + v_3 \cos 2w + \cdots + v_n \cos{(n-1)w}}{T(z+tv,w)}$

$h''(t) = \frac{(v_1 + v_2\cos w + v_3 \cos 2w + \cdots + v_n \cos{(n-1)w})^2}{T(z+tv,w)^2} \geq 0$
This means that $g(x,w)$ is convex for a fixed $w$

Hence, $f(x) = \int_0^{2\pi} g(x,w)\,dw$ is like an infinite non-negative weighted sum. Hence, $f(x)$ is convex!

---